The Annals of Mathematical Statistics
1971, Vol. 42, No. 2, 607–615

THE LAW OF THE ITERATED LOGARITHM FOR EMPIRICAL DISTRIBUTIONS[1]

BY HELEN FINKELSTEIN

University of California, Berkeley

A law of the iterated logarithm is derived for the empirical distribution functions of a sequence of independent identically distributed random variables. Convergence is in the uniform topology on the space of functions on the reals with discontinuities of the first kind only. The proof depends on a law of the iterated logarithm for independent identically distributed vector-valued random variables.

1. Introduction. Let $X_1, X_2, \cdots$ be independent identically distributed random variables defined on a probability space $(\Omega, \mathscr{F}, P)$ with distribution function $F(x)$ defined in an interval $[a, b]$. Let $\mathscr{E} = \mathscr{E}[a, b]$ be the space of functions on $[a, b]$ with the norm $\sup_x |f(x)|$ for $f \in \mathscr{E}$ and distance $(f, g) = \sup_x |f(x) - g(x)|$ for $f \in \mathscr{E}$ and $g \in \mathscr{E}$.

Suppose X_1 has finite expectation EX_1 and finite variance V^2. Let S_n be the function in $\mathscr{E}$ defined as follows: $S_n(i/n) = \sum_{k=1}^{i}[(X_k - EX_1)/V]$ for $i = 0, 1, \cdots, n$, and S_n is linear in the intervals $[i-1/n, i/n]$ for $i = 1, 2, \cdots, n$. Then Donsker's Theorem states

$$\frac{1}{n^{\frac{1}{2}}} S_n \to B \quad \text{in distribution} \tag{1}$$

where B is standard Brownian motion in $\mathscr{E}$. Strassen proved in [4] that with probability 1 the sequence $[S_n/(2n \log \log n)^{\frac{1}{2}}]_{n=3,4,\ldots}$ is relatively compact in $\mathscr{E}$ and the set of its limit points is the set of functions f in $\mathscr{E}$ such that

(i) $f(0) = 0$,

(ii) f is absolutely continuous with respect to Lebesgue measure, and

(iii) $\int_0^1 (f')^2 \leqq 1$

where f' is the derivative of f determined a.e. with respect to Lebesgue measure.

For $\omega \in \Omega$ and $x \in [0, 1]$ let $F_n(\omega, x)$ be the empirical distribution of the X_i at stage n; that is, $nF_n(\omega, x)$ is the number of $X_1(\omega), X_2(\omega), \cdots, X_n(\omega)$ which are less than x.

Let $\mathscr{D}$ be the space of functions on $[0, 1]$ which are right continuous and have left limits everywhere. Give $\mathscr{D}$ the Skorohod topology: let the distance between two elements, f and g, of $\mathscr{D}$ be

$$\inf_{\lambda \in \wedge}(\|\lambda(f) - \lambda(g)\| + \|\lambda - \varepsilon\|)$$

where $\wedge$ is the set of strictly increasing continuous mappings of $[0, 1]$ onto itself and $\varepsilon \in \wedge$ is the identity map. Then if $X_1, X_2, \cdots$ have the uniform distribution

Received July 28, 1970.

[1] This research was partially supported by research grant No. GM 10525 from the National Institutes of Health, Public Health Service.

on [0, 1] the analog to (1) in the case of empirical distributions is $n^{\frac{1}{2}}(F_n(\omega, x) - nx) \to \beta(x)$ in distribution where $\beta(x)$ is Brownian motion tied to 0 at time 1.

It is reasonable, then, to expect an analog to Strassen's theorem in the case of empirical distributions. Theorem 1 provides this result for variables with the uniform distribution on [0, 1]. Theorem 2 is a generalization of Theorem 1 to variables with an arbitrary distribution function.

2. The law of the iterated logarithm for empirical distributions. Let $X_1, X_2, \cdots$ have the uniform distribution on [0, 1]. Let

$$G_n(\omega, x) = \frac{nF_n(\omega, x) - nF(x)}{[2n \log \log n]^{\frac{1}{2}}}.$$

Let K be the set of elements f of $\mathscr{E}[0, 1]$ such that

(i) $f(0) = f(1) = 0$,

(ii) f is absolutely continuous with respect to Lebesgue measure, and

(iii) $\int_0^1 (f')^2 \leqq 1$

where f' is the derivative of f determined a.e. with respect to Lebesgue measure.

THEOREM 1. *There is a set $\Omega_0 \in \mathscr{F}$ such that $P(\Omega_0) = 1$ and for all $\omega \in \Omega_0$ the sequence $(G_n(\omega, \cdot))_{n=3, 4, \cdots}$ is relatively compact in $\mathscr{E}$ and the set of its limit points is K.*

The compactness of K is a consequence of the following lemma, due to Riesz ([2] page 75).

LEMMA 1. *Let f be a real-valued function on the unit interval. The following two conditions are equivalent:*

1. *f is absolutely continuous with respect to Lebesgue measure and*

$$\int_0^1 (f')^2 \leqq c < \infty.$$

2. $\displaystyle\sum_{i=1}^{s} \frac{(f(x_i) - f(x_{i-1}))^2}{x_i - x_{i-1}} \leqq c$ *for every finite partition $\{x_0, x_1, \cdots, x_s\}$ of* [0, 1].

Chung ([1]) and Smirnov ([3]) proved independently that if the distribution function of X_i is continuous

$$\limsup_n (\sup_x G_n(\cdot, x)) \leqq \tfrac{1}{4} \text{ a.e.} \tag{2}$$

The anomalous factor $\frac{1}{4}$ is explained by the following property of K, which follows from Lemma 1:

$$(f(s+t) - f(s))^2 \leqq t(1-t) \leqq \tfrac{1}{4} \tag{3}$$

for $0 \leqq s \leqq s+t \leqq 1$ and for all $f \in K$.

PROOF. Rearrange the function f as follows: Let

$$\begin{aligned} g(x) &= f(x+s)-f(s) && \text{for } 0 \leqq x \leqq t, \\ &= f(x-t)+f(s+t)-f(s) && \text{for } t \leqq x \leqq s+t, \\ &= f(x) && \text{for } s+t \leqq x \leqq 1. \end{aligned}$$

Then g is also an element of K. Applying Lemma 1 to g yields $(g(t))^2 \leqq t(1-t)$. But $g(t) = f(x+t)-f(t)$. □

PROOF OF THEOREM 1. The proof of Theorem 1 depends on the bound (2) and on a generalization of the Law of the Iterated Logarithm for sums of independent real-valued random variables (Lemma 2).

LEMMA 2. *Let $\mathbf{Z}_1, \mathbf{Z}_2, \cdots$ be independent identically distributed random vectors with values in m-dimensional Euclidean space R^m, with*

$$E\mathbf{Z}_1 = \mathbf{0}$$

$$(4) \qquad E\mathbf{Z}_1\mathbf{Z}_1^T = I^m \quad (\textit{the m-dimensional identity matrix}).$$

Let

$$\mathbf{\Sigma}_n = \frac{\sum_{i=1}^n \mathbf{Z}_i}{(2n \log \log n)^{\frac{1}{2}}}.$$

Then with probability 1 *the sequence $(\mathbf{\Sigma}_n)_{n=3,4,\cdots}$ is relatively compact and the set of its limit points is*

$$B_m = \{x \in R^m : ||x|| \leqq 1\}$$

where $||\cdot||$ is the Euclidean norm in R^m.

PROOF OF LEMMA 2. Lemma 2 is true if $m = 1$, i.e. if the $\mathbf{Z}_i$ are real-valued. If T is a bounded linear functional on R^m, then $T\mathbf{Z}_1, T\mathbf{Z}_2, \cdots$ are independent identically distributed random variables with $E(T\mathbf{Z}_1) = 0$ and $E(T\mathbf{Z}_1)^2 = ||T||$ by (4), so

$$(5) \qquad \limsup_n T\mathbf{\Sigma}_n = ||T|| \text{ a.e.}$$

Since the conjugate space $\hat{R}^m = R^m$ of R^m is separable, $\limsup_n T\mathbf{\Sigma}_n = ||T||$ for all $T \in \hat{R}^m$ with probability 1. Choose a point $\omega \in \Omega$ such that $\limsup_n T\mathbf{\Sigma}_n(\omega) = ||T||$ for all $T \in \hat{R}^m$. Then $\limsup_n ||\mathbf{\Sigma}_n(\omega)|| \geqq 1$.

Suppose $\limsup_n ||\mathbf{\Sigma}_n(\omega)|| = 1+\eta$. There is a sequence of functionals $\{T_n\}$, with $||T_n|| = 1$ for all n, such that $\limsup_n T_n \mathbf{\Sigma}_n = 1+\eta$. Since $\{T \in \hat{R}^m : ||T|| = 1\}$ is compact the sequence $\{T_n\}$ has a limit point L. $||L|| = 1$ and $\limsup_n L\mathbf{\Sigma}_n(\omega) = 1+\eta$. Then by (5) $\eta = 0$, so

$$(6) \qquad \limsup_n ||\mathbf{\Sigma}_n|| = 1 \quad \text{with probability} \quad 1.$$

Let $S_m = \{z \in R^m : ||z|| = 1\}$. Suppose $z_0 \in S_m$ and let $T_0 x = \langle x, z_0\rangle$ for $x \in R^m$ where $\langle x, y\rangle = \sum_{i=1}^m x_i y_i$ for $x = (x_1, \cdots, x_m)$ and $y = (y_1, \cdots, y_m)$ in R^m.

$$\limsup_n \langle \mathbf{\Sigma}_n, z_0\rangle = 1 \text{ a.s.} \tag{7}$$

Let $x \varepsilon R^m$, and for $0 \leqq \delta \leqq 1$ let $||x|| \leqq 1+\delta$ and $\langle x, z_0\rangle \geqq 1-\delta$. Then

$$||x-z_0||^2 = ||x||^2+||z_0||^2-2\langle x, z_0\rangle \leqq 5\delta. \tag{8}$$

Then (6), (7) and (8) imply that z_0 is a limit point of $(\mathbf{\Sigma}_n)_{n=3,4,\cdots}$ with probability 1. Therefore with probability 1

(9) the set of limit points of $(\mathbf{\Sigma}_n)_{n=3,4\cdots}$ contains S_m.

Let π project R^{m+1} onto R^m as follows: $\pi(x_1, \cdots, x_m, x_{m+1}) = (x_1, \cdots, x_m)$ for $(x_1, \cdots, x_{m+1}) \in R^{m+1}$.
Let $Y_1, Y_2, \cdots$ be independent identically distributed variables with mean 0, variance 1, and which are independent of the $\mathbf{Z}_i$ for $i = 1, 2, \cdots$.
Let $\mathbf{Z}_i' = (\mathbf{Z}_i, Y_i)$. Then $\pi\mathbf{Z}_i' = Z_i$. Let

$$\mathbf{\Sigma}_n' = \frac{\sum_{i=1}^n \mathbf{Z}_i'}{[2n \log\log n]^{\frac{1}{2}}}$$

$$\Sigma_n = \pi\mathbf{\Sigma}_n'.$$

Now (9) is true for all m; in particular, with probability 1 the set of limit points of the sequence $(\mathbf{\Sigma}_n')_{n=3,4,\cdots}$ contains S_{m+1}. Then with probability 1 the set of limit points of $(\mathbf{\Sigma}_n)_{n=3,4,\cdots}$ contains $\pi(S_{m+1}) = B_m$. Since (6) implies that the set of limit points of $(\mathbf{\Sigma}_n)_{n=3,4,\cdots}$ is contained in B_m with probability 1, Lemma 2 follows. □

Choose a large integer m, and divide [0, 1] into m equal subintervals $I_i = [i+1/m, i/m]$ for $i = 1, 2, \cdots, m$. For each $n = 1, 2, \cdots$ and $i = 1, 2, \cdots, m$ define

$$\begin{aligned} Y_{ni} &= 1 \quad \text{if } X_n \in I_i; \\ &= 0 \quad \text{if } X_n \notin I_i. \end{aligned}$$

The vectors $\mathbf{Y}_n = (Y_{n1}, Y_{n2}, \cdots, Y_{nm})$ are independent identically distributed random elements of R^m with

$$E\mathbf{Y}_1 = \left(\frac{1}{m}, \cdots, \frac{1}{m}\right), \quad E(\mathbf{Y}_1 - E\mathbf{Y}_1)(\mathbf{Y}_1 - E\mathbf{Y}_1)^T = \Gamma$$

where Γ is the matrix $(\gamma_{ij})_{i=1,\dots,m,\ j=1,\dots,m}$ and

$$\begin{aligned} \gamma_{ij} &= \frac{1}{m} - \frac{1}{m^2} \quad \text{if } i = j; \\ &= -\frac{1}{m^2} \quad \text{if } i \neq j. \end{aligned}$$

LEMMA 3. *With probability* 1 *the sequence*

$$\left(\frac{\sum_{i=1}^{n}(\mathbf{Y}_i - E\mathbf{Y}_i)}{[2n \log\log n]^{\frac{1}{2}}}\right)_{n=3,4,\dots}$$

is relatively compact and the set of its limit points is

$$C_m = \{x \in R^m,\ x = (x_1, \cdots, x_m) : \Sigma x_i = 0 \text{ and } \Sigma x_i^2 \leqq (1/m)\}.$$

PROOF OF LEMMA 3. The range of $\mathbf{Y}_1 - E\mathbf{Y}_1$ is the hyperplane $\mathscr{H}$ defined by $\sum_{i=1}^{m} x_i = 0$ for $x = (x_1, \cdots, x_m) \in R^m$. There exist a linear transformation T from R^{m-1} to $\mathscr{H}$ and independent identically distributed random vectors $\mathbf{Z}_1, \mathbf{Z}_2, \cdots$ in R^{m-1} with $E\mathbf{Z}_1 = \mathbf{0}$ and $E\mathbf{Z}_1\mathbf{Z}_1^T = I^{m-1}$ such that

$$\mathbf{Y}_i - E\mathbf{Y}_i = T\mathbf{Z}_i$$

for $i = 1, 2, \cdots$.

For any vector $\mathbf{a} \in \mathscr{H}$, $\mathbf{a}\Gamma = (m^{-1})\mathbf{a}$, so the transformation T can be chosen to be the composition of an isometry with multiplication by the scalar m^{-1}. Then $T(B_{m-1}) = C_m$. Lemma 2 can be applied to the $\mathbf{Z}_i$, and the application of T to the result yields Lemma 3. □

Let $H_{n,m}(\cdot, x)$ be the linear interpolation of $G_n(\cdot, x)$ between the points $x = i/m$ for $i = 0, 1, \cdots, m$. That is $H_{n,m}(\omega, x) = G_n(\omega, x)$ when $x = i/m$ for $i = 0, 1, \cdots, m$. $H_{n,m}(\omega, x)$ is linear in $I_i = [(i-1)/m, i/m]$ for $i = 1, \cdots, m$.

LEMMA 4. *There is a set* Ω_1 *in* $\mathscr{F}$ *such that* $P(\Omega_1) = 1$ *and for all* $\omega \in \Omega_1$ *for all fixed* m *as* $n \to \infty$ *the sequence* $(H_{n,m}(\omega, \cdot))_{n=3,4,\cdots}$ *is relatively compact in* $\mathscr{E}$ *and the set of its limit points is*

$$J_m = \{f \in K : f \text{ is linear in } I_i \text{ for } i = 1, 2, \cdots, m\}.$$

PROOF OF LEMMA 4. First, observe that

$$H_{n,m}\left(\frac{i}{m}\right) - H_{n,m}\left(\frac{i-1}{m}\right) = \frac{\sum_{k=1}^{n}(Y_{ki} - EY_{ki})}{[2n \log\log n]^{\frac{1}{2}}} \qquad \text{for } i = 1, 2, \cdots, m. \tag{10}$$

Let $\mathscr{L}_m$ be the space of continuous functions on [0, 1] which are 0 at 0 and linear in the intervals I_i for $i = 1, 2, \cdots, m$. Give $\mathscr{L}_m$ the uniform topology. So $H_{n,m}(\omega, \cdot)$ is an element of $\mathscr{L}_m$.

Let V be the mapping from $\mathscr{L}_m$ to R^m which maps $f \in \mathscr{L}_m$ into the vector $(f(i/m) - f(i-1)/m))_{i=1,2,\dots m}$. Then (10) and Lemma 3 imply that there is a set Ω_{1m} in $\mathscr{F}$ such that $P(\Omega_{1m}) = 1$ and for all $\omega \in \Omega_{1m}$ the sequence

$$(VH_{n,m}(\omega, \cdot))_{n=3,4,\dots}$$

is relatively compact in R^m and the set of its limit points is C_m. Since V is $1-1$ and bicontinuous, for all $\omega \in \Omega_{1m}$ the sequence $(H_{n,m}(\omega, \cdot))_{n=3,4,\cdots}$ is relatively compact in $\mathscr{E}$ and the set of its limit points is $V^{-1}(C_m) = J_m$. Then $\Omega_1 = \bigcap_{m=1}^{\infty} \Omega_{1m}$. □

Let $f \in K$ and let $g \in J_m$ be its linear approximation: $g(x_i) = f(x_i)$ for $i = 0, 1, \cdots, m$. It follows from (3) that for $x \in I_i$

$$g(x) - f(x) \leqq |f(x_i) - f(x_{i-1})| + |f(x) - f(x_{i-1})|$$
$$\leqq 2/m^{\frac{1}{2}}.$$

Then J_m is a good approximation to K in the sense that for all $f \in K$ there exists a $g \in J_m$ such that $\|f - g\| \leqq 2/m^{\frac{1}{2}}$. So Lemma 4 implies

COROLLARY 1. *For all* $\omega \in \Omega_1$

(i) *the sequence* $(H_{n,m}(\omega, \cdot))_{n=3,4,\cdots}$ *is relatively compact,*

(ii) *the set of its limit points is contained in K, and*

(iii) *for all* $f \in K$ *and all* $\varepsilon > 0$, *if* $m > \frac{1}{2}\varepsilon^2$ *then* $\sup_{x \in [0,1]} |H_{n,m}(\omega, x) - f(x)| < \varepsilon$ *infinitely often as* $n \to \infty$.

LEMMA 5. *For each integer m and each* $i = 1, 2, \cdots, m$ *there is a set* Ω_{mi} *in* $\mathscr{F}$ *with* $P(\Omega_{mi}) = 1$ *and for each* $\omega \in \Omega_{mi}$ *there is a positive integer* $N(\omega)$ *such that if* $n > N(\omega)$

$$\sup_{x \in I_i} |H_{n,m}(\omega, x) - G_n(\omega, x)| < 1/m^{\frac{1}{2}}.$$

PROOF OF LEMMA 5. Fix m and i. Let

$$d(n, \omega) = [2n \log \log n]^{\frac{1}{2}} \sup_{x \in I_i} |H_{n,m}(\omega, x) - G_n(\omega, x)|.$$

Let $v(0) = 0$

$$v(n) = nF_n\left(\cdot, \frac{i}{m}\right) - nF_n\left(\cdot, \frac{i-1}{m}\right) \qquad \text{for } n = 1, 2, \cdots$$

$v(n)$ is the number of $X_1, \cdots, X_n$ which fall in I_i. For all $n = 1, 2, \cdots$

$$\text{(11)} \qquad \begin{aligned} v(n) &= v(n-1) + 1 && \text{if } X_n \in I_i; \\ &= v(n-1) && \text{if } X_n \notin I_i. \end{aligned}$$

By algebra

$$\text{(12)} \qquad d(n, \cdot) = \sup_{x \in I_i} \left| nF_n(\cdot, x) - nF_n\left(\cdot, \frac{i-1}{m}\right) - v(n)m\left(x - \frac{i-1}{m}\right)\right|.$$

It can be seen from (12) that d_n depends only on those X_i out of $X_1, \cdots, X_n$ which fall in I_i.

Let $k \leqq j_1 < j_2 < \cdots < j_p \leqq K$ be positive integers. Let A be the set in $\mathscr{F}$ on which out of $X_k, X_{k+1}, \cdots, X_K$ all of $X_{j_1}, X_{j_2}, \cdots, X_{j_p}$, but no others, take values in I_i. The joint conditional distribution of $X_{j_1}, X_{j_2}, \cdots, X_{j_p}$ given A is just the distribution of p independent random variables with the uniform distribution on I_i.

So in order to examine the distribution of d, define X_1^*, X_2^*, $\cdots$ independent random variables with the uniform distribution on I_i. Let $F_n^*(\omega, x)$ be the empirical distribution of the X_i^* at stage n. Set

$$d^*(n, \cdot) = \sup_{x \in I_i} \left|nF_n^*(\cdot, x) - nm\left(x - \frac{i-1}{m}\right)\right|.$$

Let $\lambda(z) = [2z \log\log z]^{\frac{1}{2}}$.

If $(s_k, s_{k+1}, \cdots, s_K)$ is a sequence of possible values of $(v(k), v(k+1), \cdots, v(K))$ according to (11), the set

$$B = \{v(k) = s_k,\ v(k+1) = s_{k+1}, \cdots, v(K) = s_K\}$$

is of the form of the set A above with $p = s_K - s_k$. Then the conditional probability

$$\begin{aligned} &P\{d(n,\cdot) > \lambda(v(n)) \quad \text{for some } n : k \leqq n \leqq K \mid B\} \\ &\quad = P\{d^*(s,\cdot) > \lambda(s) \quad \text{for some } s_k \leqq s \leqq s_K\}. \end{aligned} \tag{13}$$

Let S be the set of all possible sequences of values of $v(n)$. For $\mathbf{s} \in S$ let $\mathbf{s} = (s_i)_{i=1,2,\ldots}$. Then (13) implies

$$\begin{aligned} &P\{d(n,\cdot) > \lambda(v(n)) \quad \text{for some } n : k \leqq n \leqq K\} \\ &\quad = \int_S P\{d^*(s,\cdot) > \lambda(s) \quad \text{for some } s : s_k \leqq s \leqq s_K\} P_v(d\mathbf{s}) \end{aligned} \tag{14}$$

where P_v is the distribution of $(v(n))_{n=1,2,\ldots}$.

Let $C = \{\mathbf{s} \in S:\ \lim_n s_n = \infty\}$. By the Strong Law of Large Numbers, $P_v(C) = 1$. Then application of the Monotone Convergence theorem to (14) letting $K \uparrow \infty$ then $k \uparrow \infty$ yields

$$\begin{aligned} &P(\limsup_n \{d_n > \lambda(v(n))\}) \\ &\quad = \int_C P(\limsup_{s_n} \{d^*(s_n, \cdot) > \lambda(s_n)\}) P_v(d\mathbf{s}) \\ &\quad = \int_C P(\limsup_s \{d^*(s,\cdot) > \lambda(s)\}) P_v(d\mathbf{s}). \end{aligned}$$

But (2) implies that $P(\limsup_s \{d^*(s, \cdot) > \lambda(s)\}) = 0$; therefore

$$P(\limsup_n \{d(n, \cdot) > \lambda(v(n))\}) = 0. \tag{15}$$

The Strong Law of Large Numbers implies that

$$\frac{\lambda(v(n))}{\lambda(n)} \to \frac{1}{m^{\frac{1}{2}}} \text{ a.s.} \qquad \text{as } n \to \infty. \tag{16}$$

Then since

$$\sup_{x \in I_i} \left|H_{n,m}(\cdot, x) - G_n(\cdot, x)\right| = \frac{d(n, \cdot)}{\lambda(n)},$$

(15) and (16) imply Lemma 5. □

To prove Theorem 1, consider the set

$$\Omega_0 = \Omega_1 \bigcap \left(\bigcap_{m=1}^{\infty} \bigcap_{i=1}^{m} \Omega_{im} \right).$$

$$P(\Omega_0) = 1.$$

It follows from Lemma 5 and the corollary to Lemma 4 that for all $\omega \in \Omega_0$

(i) the sequence $(G_n(\omega, \cdot))_{n=1,2,\ldots}$ is relatively compact,

(ii) the set of its limit points is contained in K, and

(iii) for all $f \in K$ and all $\varepsilon > 0$, $\sup_{x\in[0,1]} |G_n(\omega, x) - f(x)| < \varepsilon$ infinitely often as $n \to \infty$. ▯

3. An application of Theorem 1. Let K be the set of functions $f \in \mathscr{D}$ such that

(i) $f(0) = f(1) = 0$,

(ii) f is absolutely continuous with respect to Lebesgue measure, and

(iii) $\int_0^1 (f')^2 \leqq 1$.

Then

$$\sup_{f\in K}\left(\int_0^1 f^2\right) = \frac{1}{\pi^2}\,. \tag{17}$$

The extreme value is attained by $f(x) = (2^{\frac{1}{2}}/\pi) \sin(\pi x)$.

PROOF. Since K is uniformly compact there is a function, say h, for which the sup is attained.

By Calculus of variations there is a constant λ (a Lagrange multiplier) such that

$$\int_0^1 h(x) f(x)\, dx + \lambda \int_0^1 h'(x) f'(x) = 0 \tag{18}$$

for all functions $f \in \mathscr{D}$ satisfying (i) and (ii).

Let $f(x) = \sin(\pi x)$ in (18). Integrating by parts

$$\int_0^1 h'(x) \cos(\pi x)\, dx = \pi \int_0^1 h(x) \sin(\pi x)\, dx. \tag{19}$$

The function h can be chosen so that the right-hand integral is not zero. Then (18) and (19) yield

$$\lambda = \frac{-\int_0^1 h(x) \sin(\pi x)\, dx}{\pi^2 \int_0^1 h(x) \sin(\pi x)\, dx} = \frac{-1}{\pi^2}\,. \tag{20}$$

If h maximizes $\int_0^1 f^2$ in K, then $\int_0^1 (h')^2 = 1$. Setting $f(x) = h(x)$ in (18) and using (20) we get

$$\int_0^1 [h(x)]^2 = -\lambda = \frac{1}{\pi^2}\,. \quad ▯$$

The bound (17) implies the following corollary to Theorem 1:

COROLLARY. *With probability* 1

$$\limsup \frac{\int_0^1 [F_n(\cdot, x) - nx]^2\, dx}{2n \log\log n} = \frac{1}{\pi^2}.$$

4. A generalization of Theorem 1. Let $X_1, X_2, \cdots$ have a continuous distribution function $F(x)$ defined on an interval $[a, b]$. Let K_F be the set of functions $f \in \mathscr{C}[a, b]$ such that

(i) $f(a) = f(b) = 0$,

(ii) f is absolutely continuous with respect to F, and

(iii) $\int_a^b (df/dF)^2\, dF \leqq 1$ where (df/DF) is the derivative of f with respect to F defined a.e. with respect to F.

THEOREM 2. *There is a set* $\Omega_F \in \mathscr{F}$ *such that* $P(\Omega_F) = 1$ *and for all* $\omega \in \Omega_F$ *the sequence* $(G_n(\omega, \cdot))_{n=3,4,\ldots}$ *is relatively compact in* $\mathscr{C}[a, b]$ *and the set of its limit points is* K_F.

PROOF OF THEOREM 2. Theorem 2 can be proved using the arguments used in the proof of Theorem 1 with a few changes. For example the intervals I_i should be redefined as follows: let $x_0, x_1, \cdots, x_m$ be points in $[a, b]$ such that $F([x_{i-1}, x_i]) = 1/m$ for $i = 1, 2, \cdots, m$. Let $I_i = [x_{i-1}, x_i]$ for $i = 1, 2, \cdots, m$. The functions $H_{n,m}$ must also be redefined as the interpolation of G_n between the points x_i according to F; that is,

$$H_{n,m}(\cdot, x_i) = G_n(\cdot, x_i) \qquad \text{for } i = 0, 1, \cdots, m.$$

For $x \in I_i$, $i = 1, 2, \cdots, n$,

$$H_{n,m}(\cdot, x) = G_n(\cdot, x_{i-1}) + m(F(x) - F(x_{i-1}))(G_n(x_i) - G_n(x_{i-1})).$$

The remaining changes should be obvious.

Acknowledgments. I should like to thank David Freedman for his guidance and encouragement, Peter Bickel for his interesting comments, and Volker Strassen for suggesting the investigation of this subject.

REFERENCES

[1] CHUNG, K.-L. (1949). An estimate concerning the Kolmogoroff limit distribution. *Trans. Amer. Math. Soc.* **67** 36–50.

[2] RIESZ, F. and SZ.-NAGY, B. (1955). *Functional Analysis*, Frederick Unger, New York.

[3] SMIRNOFF, N. (1939). Sur les écarts de la courbe de distribution empirique. *Mat. Sbornik.* **6** 3–26. (Russian-French summary).

[4] STRASSEN, V. (1964). An invariance principle for the Law of the Iterated Logarithm. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **3** 211–226.